TECHNICAL NOTE Nº IDB-TN-3199

Generative AI in Education

A Framework for Leveraging Digital Tools in Latin American Classrooms

Eduardo Levy Yeyati
Virginia Robano
Emiliano Pereiro
Camila Porto
Víctor Koleszar

Inter-American Development Bank
Department of Research and Chief Economist

December 2025

Generative AI in Education:

A Framework for Leveraging Digital Tools in Latin American Classrooms

Eduardo Levy Yeyati*
Virginia Robano**
Emiliano Pereiro**
Camila Porto**
Víctor Koleszar**

* The Brookings Institution and Universidad Torcuato Di Tella
** Centro Ceibal

Inter-American Development Bank
Department of Research and Chief Economist

December 2025

Cataloging-in-Publication data provided by the
Inter-American Development Bank
Felipe Herrera Library

Generative AI in education: a framework for leveraging digital tools in Latin American classrooms / Eduardo Levy Yeyati, Virginia Robano, Emiliano Pereiro, Camila Porto, Víctor Koleszar.
p. cm. — (IDB Technical Note ; 3199)
Includes bibliographical references.
1. Artificial intelligence-Educational applications-Latin America. 2. Education-Effect of technological innovations on-Latin America. 3. Generative artificial intelligence-Social aspects-Latin America. I. Levy Yeyati, Eduardo. II. Robano, Virginia. III. Pereiro, Emiliano. IV. Porto, Camila. V. Koleszar, Víctor. VI. Inter-American Development Bank. Department of Research and Chief Economist. VII. Series.
IDB-TN-3199

http://www.iadb.org

Abstract[1]

Generative AI (GenAI) is entering classrooms with both promise and uncertainty. This paper develops a framework that organizes instructional uses of GenAI into four functions: content preparation, explanation, practice and feedback, and motivation and exploration.Using evidence from Uruguay's Ceibal program, including platform data and observation from classroom sessions in a computational- thinking program, we show that patterns of use are not neutral. Female teachers and students engage more with tools for preparation and structured practice, while their male counterparts dominate exploratory platforms. These trajectories risk reinforcing inequality by channeling GenAI toward automation for some and augmentation for others. To test an alternative path, we design a randomized controlled trial that integrates a chatbot for explanation and feedback, explicitly structured for teacher mediation. The paper concludes with implications for Latin American education systems, and highlights that deliberate design and phased implementation are essential to ensure GenAI heightens inclusion rather than disparity.

JEL Classification: I21, J16, O33, J24, D63

Keywords: artificial intelligence, education, ChatGPT, complementarity, LLM, automated tutor, chatbot, classroom, teaching, computational thinking, economics of gender, nonlabor discrimination, occupational choice, equity, justice, analysis of education, technological change

[1] We thank Daron Acemoglu, Elena Arias Ortiz, Julian Cristia, and Gregory Elacqua for their excellent comments. We also gratefully acknowledge the valuable support of Elton Mancilla in producing this document. This work was financed with support from the Latin America and the Caribbean Research Network of the Inter-American Development Bank. The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the Inter-American Development Bank, its Board of Directors, or the countries they represent. Errors are our responsibility.

Introduction

As generative artificial intelligence (GenAI) technologies rapidly advance, their potential to transform education has become both promising and contentious. Unlike earlier waves of educational technology (such as learning management systems or static content repositories), GenAI introduces dynamic, dialogic tools that can create, adapt, and interact in real time with learners (Mollick and Mollick, 2023). This capacity to generate content, scaffold learning, and engage students through conversational interfaces positions GenAI, and particularly chatbots, as a novel layer of educational support. However, realizing this potential requires careful attention to how GenAI tools are integrated into classroom practices, especially in regions where educational systems face structural challenges. Historical experience shows that technology does not automatically deliver broad-based prosperity; progress depends on how power and incentives shape the use of technology (Acemoglu and Johnson, 2023). Technological breakthroughs may fuel long-run economic growth, but only if accompanied by complementary investments in skills and institutions (Aghion, Jones, and Jones, 2017).

Latin America exemplifies this challenge . While countries in the region have made strides in expanding digital access through initiatives like Ceibal in Uruguay since 2007 (Robano, 2023, 2024; Robano, González and Galland, 2025; Curi et al., 2020), persistent inequalities in teacher training, resource allocation, and gender-based participation hinder equitable adoption of new technologies (Carvajal, Franco and Isaksson, 2024; Arias-Ortiz et al., 2024). At the same time, teacher shortages, high turnover, and learning gaps (particularly in mathematics and science) underscore the urgent need for scalable solutions that support classroom instruction (Saavedra et al., 2024; INEEd, 2024, Castro et al., 2024). In this context, chatbots emerge as a potential tool to complement human teachers by providing real-time explanations, addressing routine queries, and supporting differentiated instruction (Brynjolfsson et al., 2024; Capdehourat et al., 2025).

Students and teachers are already relying on these systems, often in the absence of clear guidance. This reality makes it imperative to identify which applications genuinely enhance learning and which may undermine it, so that schools and governments can steer adoption toward effective and equitable outcomes. As Joubin (2024) argues, ensuring the trustworthiness of GenAI in educational contexts requires not only technical safeguards but also pedagogical strategies that foster critical engagement, particularly in disciplines in which interpretive reasoning is central. Emerging evidence suggests that while chatbots can enhance learning engagement and provide

personalized feedback, they may also encourage overreliance on the technology itself, reduce critical thinking, and reinforce existing inequalities if not properly designed and integrated (Bastani et al., 2025; Zhai et al., 2024). Gender disparities are a particular concern: Girls and female teachers often show lower engagement with GenAI tools, reflecting both structural barriers and self-perception biases (Robano, González and Galland, 2025; Porto et al., 2024, Carvajal, Franco and Isaakson, 2024). Additionally, students may use chatbots in ways that substitute for rather than enhance their own learning efforts, raising concerns about cognitive offloading and diminished effort (Lehmann, Cornelius and Sting, 2025).

This paper addresses the question: **How can chatbots be integrated into classrooms so that they complement (rather than substitute for) teachers while advancing equity and learning effectiveness?** Here, we conceptualize complementarity as a continuum, ranging from light-touch assistance in high-capacity schools to more autonomous support where teacher availability is limited. We focus on Latin American classrooms as a case study, drawing on evidence from Uruguay's Ceibal Computational Thinking and Artificial Intelligence Program (CT/AI-P) and Ceibal's gender dashboard, a monitoring tool designed to quantify gender gaps in technology across the dimensions of access, use, and benefits (Robano, 2023). T his case offers granular usage data and a policy environment open to experimentation.

Latin American classrooms operate under significant structural constraints. Despite the expansion of educational coverage, low student attendance persists as a pressing challenge. In Uruguay, for instance, the average school attendance rate fell to 86.1 percent in 2023, with 59.3 percent of students chronically absent (INEEd, 2024). This reality forms a critical boundary condition for our study: The potential for any educational technology, including GenAI, is limited when students are physically absent. Our design must therefore account for supporting catch-up and maintaining engagement amid these challenges.

We argue that the potential of GenAI in education depends less on the technology itself than on how it is adopted. Structured, student-facing applications, in which learners interact with tools under predefined pedagogical constraints and active teacher mediation, offer a viable path to leverage benefits while mitigating risks. Our contribution is threefold: (1) We develop an economic framework that maps GenAI capabilities to core teaching tasks, analyzing each through the lens of complementarity and automation; (2) we synthesize global evidence to diagnose how preexisting patterns of technology usage create externalities that risk amplifying inequalities; and

(3) we propose a randomized controlled trial (RCT) designed not only to measure learning impacts but to test whether teacher mediation can steer GenAI toward more equitable and complementary outcomes.

By grounding our analysis in both theory and practice, we can inform policymakers, educators, and researchers seeking to harness GenAI for educational equity and quality. The findings and proposals in this paper contribute to a growing body of evidence on how GenAI can support inclusive, context-sensitive innovations in education.

The paper is structured as follows: Section 2 proposes a conceptual framework to integrate GenAI in education, reviewing the literature and proposing a set of four principles for evaluating GenAI uses in the classroom. Section 3 describes and analyzes the data, including data from Ceibal's CT/AI-P and gender dashboard. Section 4 presents the proposed methodological framework and outlines the experimental design. Sections 5 and 6 discuss the policy implications and conclude, while Section 7 suggests a research agenda.

2. A Task-Based Framework for AI in the Classroom

2.1. Moving Beyond the Hype: From AI to Instructional Functions

The discourse surrounding GenAI in education often oscillates between boundless enthusiasm and profound skepticism. To move beyond this dichotomy, we propose a pragmatic, task-based framework. Drawing on the economics of technology and labor (Acemoglu and Restrepo, 2018, 2019), we analyze GenAI not as a monolithic force but through the specific instructional functions it supports. This lens allows for a granular assessment of its potential to either automate tasks or complement human expertise, a distinction critical for predicting its impact on learning and equity (Brynjolfsson et al., 2023).

2.2. The Four Core Functions of Instructional AI

Synthesizing the literature (Mollick and Mollick, 2023; INTEF, 2022; Bastani et al., 2025), we distill GenAI applications into four core functions that map onto fundamental teacher responsibilities:

1. **Content Preparation**

   This function involves automation of logistical and design tasks, such as generating lesson plans, quizzes, examples, and prompts (Mollick and Mollick, 2023). Its value lies in efficiency gains, freeing up teacher time for higher-value interactions. The primary risk is

the de-skilling of curriculum design if this function is used as a substitute for critical engagement with pedagogical content.

2. **Explanation and Instruction**

   This function covers the augmentation of direct teaching through conversational agents that provide step-by-step guidance, clarify concepts, and answer student questions in real time (Wang et al., 2024a; Luckin et al., 2016). It promises to make instruction scalable and personalized. The paramount risks are confabulation (factual error) and the erosion of the teacher's role as the primary authority if not carefully mediated (Mollick and Mollick, 2023; Zhai et al., 2024). These pedagogical risks exemplify the broader economic dangers of automation eroding human agency and de-skilling professionals (Acemoglu, 2021).

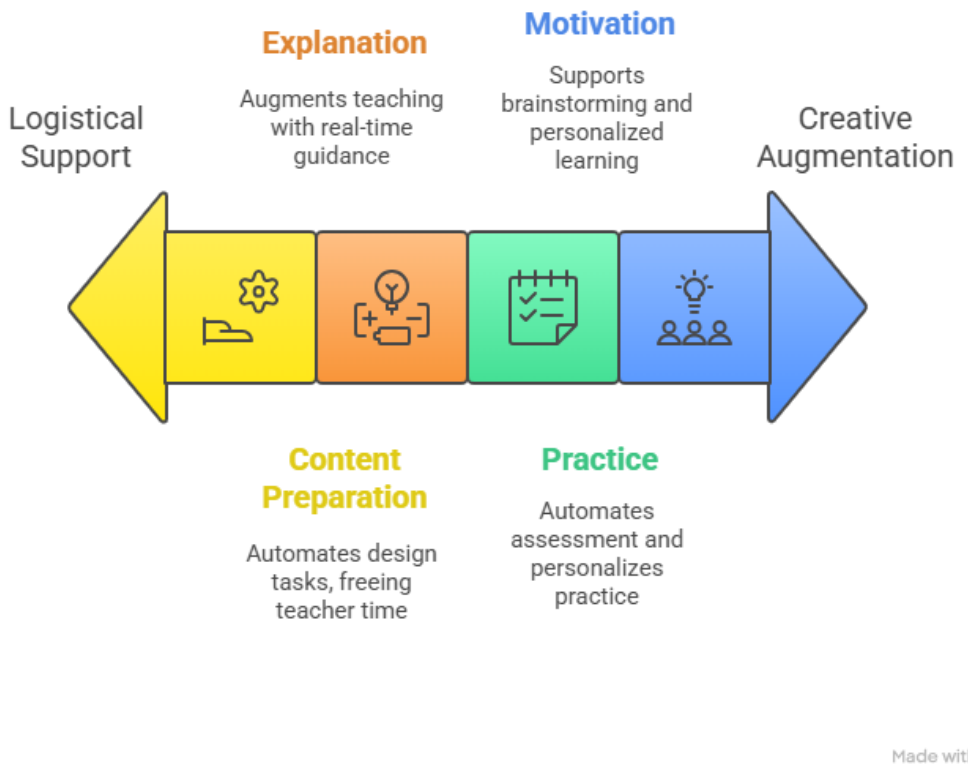
3. **Practice and Feedback**

   This function automates assessment and personalizes practice. It includes generating variations of problems, simulating scenarios, and, importantly, providing formative feedback on open-ended responses (Capdehourat et al., 2025; Henkel et al., 2024). It is distinct from explanation, as it provides opportunities for application and reinforcement rather than initial instruction. The risk lies in overly generic or inaccurate feedback and the potential for reducing critical effort (cognitive offloading) if not dosage controlled (Zhai et al., 2024; Lehmann et al., 2025).

4. **Motivation and Exploration**

   This function aims to augment student curiosity and creativity by supporting brainstorming, project ideation, and personalized learning pathways (Mazzucato and Larghi, 2024). It is the most complex and high-risk function. While it holds potential for engagement (Kim and Kwon, 2024), it can easily produce shallow or off-topic output without significant teacher scaffolding and is highly susceptible to reinforcing existing engagement gaps.

Figure 1: GenAI instructional function support



2.3. The Determinant of Impact: Teacher Mediation

International evidence consistently shows that the impact of these functions is not inherent to the technology but depends on its integration into pedagogical practice (Bastani et al., 2024; Vanzo et al., 2024; Sahito et al., 2024). The teacher's role in curating, validating, and contextualizing GenAI output is not a peripheral activity; it is the core mechanism that determines whether GenAI serves as a substitute or a complement. This aligns with economic evidence that technology benefits mid-skill workers most when it augments their capabilities rather than replacing their tasks (Brynjolfsson et al., 2023). Therefore, the critical question for policy and design is not whether to use GenAI, but how to structure its use to maximize mediation and hence complementarity.

This framework provides the necessary vocabulary to analyze current adoption patterns and design targeted interventions. The following section applies this lens to diagnose the existing

trajectory of technology use in Uruguay, revealing the preexisting conditions that will shape GenAI's equitable (or inequitable) adoption.

Table 1: Mapping of teacher tasks to GenAI: Contributions and risks

Teacher Task	GenAI Function	GenAI Tool Example	Key Contribution	Primary Risks	Our Intervention Design Choice
Planning lessons and preparing materials	Content preparation	Content generators (e.g., CREA's PowerBuddy)	Drafts outlines, quizzes, examples; saves teacher time	Outputs may be generic, erroneous, or lack local context; risk of teacher de-skilling in curriculum design	Teacher-facing support only; enabled for use in lesson preparation, but disabled during student sessions to preserve teacher agency
Delivering instruction; explaining concepts	Explanation and instruction	Conversational agents, multimodal aids	Provides step-by-step guidance; clarifies concepts; answers questions; enables scalable, personalized tutoring	Confabulation (factual errors), lack of curricular awareness, erosion of teacher's role as knowledge authority	In scope, with heavy mediation; available to students during sessions for explanation of predefined concepts, with teacher ability to review, edit, or reject all AI responses
Providing formative feedback; facilitating practice	Practice and feedback	Assessment engines, adaptive platforms	Offers immediate, limitless formative feedback; generates personalized practice problems and simulations	Generic/inaccurate feedback, cognitive offloading, reduced student effort, overreliance on repetitive practice	In scope, with bounded use; provides rubric-guided feedback on student code/answers; usage is limited by time and attempt constraints to sustain cognitive engagement

Fostering exploration and creativity; motivating students	Motivation and exploration	Generative tools for brainstorming, project ideation	Supports open-ended tasks, brainstorming, and personalized learning pathways	Superficial or off-topic output, amplification of existing engagement inequalities, lack of pedagogical structure	Out of scope for initial pilot; excluded from the current design to avoid unstructured use and focus the RCT on testing core complementary functions
This mapping underscores a central thesis: The pedagogical value of GenAI is not inherent but depends on its deliberate alignment with specific tasks and the surrounding instructional design that prioritizes teacher mediation.					

2.4. The Evidence on GenAI: Complements, Substitutes, and Externalities

The international evidence on GenAI in education, while nascent, consistently shows that its impact is not predetermined but hinges on its design and integration, validating our task-based framework. The literature reveals a clear tension between its potential as a complement to teaching and its risk as a substitute for student effort and teacher expertise, often with significant distributional consequences.

Structured implementations that enhance teacher capabilities show promise. For example, Stanford's Tutor CoPilot increased learning gains and improved the practices of less experienced tutors by providing AI-generated guidance, thus demonstrating the complementarity effect, in which AI augments mid-skill professionals (Wang et al., 2024b). Similarly, studies in Pakistan and South Korea found that AI-generated feedback improved outcomes when teachers were actively involved in interpreting and mediating the output for students (Sahito et al., 2024; Son, 2024). This aligns with broader economic evidence that technology boosts productivity most when it complements human decision-making (Brynjolfsson et al., 2023).

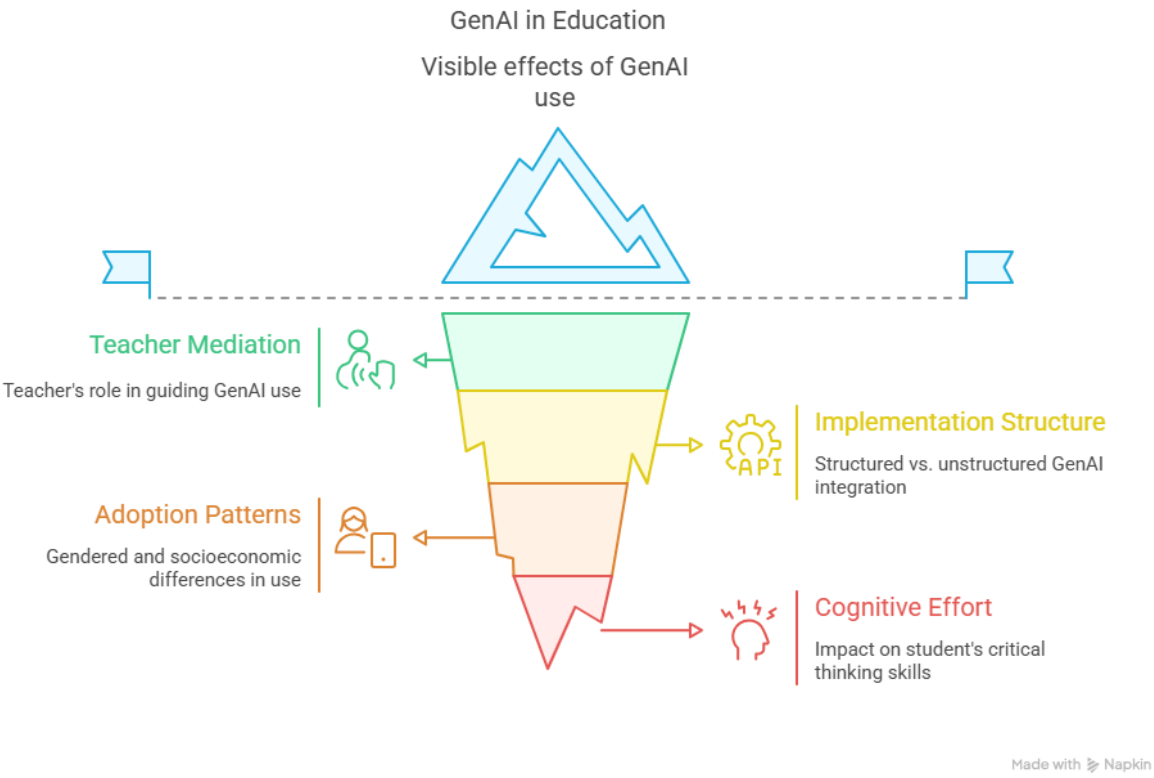
Conversely, unguided use often leads to substitution with negative results. A study in Turkey found that unstructured use of GPT tutors improved in-task performance but reduced final test scores, suggesting students used the tool as a shortcut that bypassed essential cognitive effort (Bastani et al., 2024). Other studies document increased overreliance, reduced critical thinking, and cognitive offloading when students use GenAI without pedagogical scaffolding (Zhai et al.,

2024; Lehmann et al., 2025). This risk of de-skilling and erosion of human agency is a core concern in the economics of automation (Acemoglu, 2021).

The distributional effects of these dynamics are profound. Research consistently shows that access alone is insufficient; patterns of adoption and use are gendered and socioeconomically stratified. Recent PISA data show that despite progress in infrastructure and digital access, Uruguay still faces persistent challenges in learning outcomes and educational equity (ANEP, 2023).[2] Studies from Uruguay and across Latin America confirm that girls and female teachers often show lower engagement with tech tools, especially in exploratory STEM domains, reflecting both structural barriers and self-perception biases (Robano et al., 2025; Porto et al., 2024; Carvajal et al., 2024). This is not a passive gap but an active externality: Early differences in access and confidence create feedback loops that widen disparities over time, a mechanism highlighted by Acemoglu et al. (2024) in other digital domains.

Figure 2: GenAI: More than Meets the Eye

[2] ECLAC (2024) emphasizes that addressing school dropout in Latin America requires early-detection systems and pedagogical innovation, particularly in contexts of digital exclusion and socioeconomic vulnerability, another possible area for AI tools.



The critical lesson from this literature is that teacher mediation is the key variable that determines which equilibrium prevails: complementarity or substitution. Our framework provides the structure to analyze this, and the following section applies it to the Uruguayan context to predict the default trajectory and design an intervention to alter it.

3. The Allocation of Learning Opportunities: How Platform Design and Teacher Practices Shape GenAI Readiness

The conceptual framework provides a taxonomy for GenAI's pedagogical potential. However, realizing this potential is not a technical question but an economic and institutional one. It depends critically on how the allocation of digital learning opportunities (shaped by platform design, teacher practices, and organizational routines) prepares different groups of teachers and students to engage with specific GenAI functions. This section analyzes usage patterns in Uruguay's Ceibal ecosystem through this lens, drawing on visualizations from Ceibal's gender dashboard to track disparities. We examine how the current direction of digital technology use predicts and may

constrain the future adoption of GenAI, with significant implications for equity and complementarity.

3.1. Ceibal as a Natural Laboratory for Studying Technological Externalities

Uruguay's Centro Ceibal provides a unique setting to analyze the allocation of digital opportunities. For over 18 years, it has ensured near-universal access to digital devices and connectivity, effectively holding the variable of basic access constant. This allows us to move beyond simplistic access metrics and examine the more nuanced, and ultimately more important, question of how technology is used and which groups benefit from its advanced functions.[3]

This analysis is not merely descriptive; it is diagnostic. Following Acemoglu et al. (2024), we contend that technology adoption creates externalities that shift preferences and alter equilibria. The current patterns of engagement with pre-GenAI digital tools reveal the existing trajectory and allocation mechanisms. If left unaddressed, these patterns will inevitably shape (and likely exacerbate) the adoption and impact of GenAI, determining whether it serves as a tool for complementarity or a force for inequality.

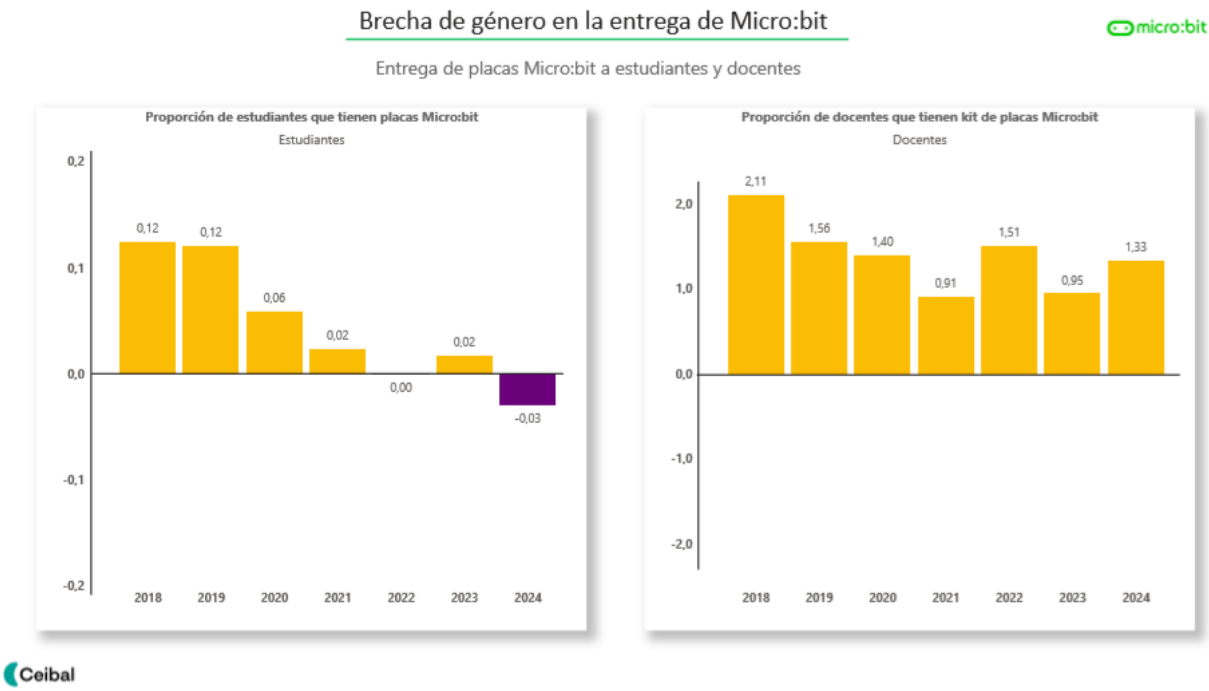
3.2. Gendered Externalities in the Allocation of Digital Tools

The first evidence of unequal allocation is found in access to specialized hardware. While laptop distribution is gender balanced, access to tools that enable the exploration function is not. Figure 3 shows that male students and teachers are significantly more likely to receive and use specialized tools like micro:bit boards.[4] This disparity is not a random outcome; it is the result of organizational routines, teacher preferences, and implicit signals about who belongs in hands-on programming and exploratory learning, creating a powerful negative externality: Early male dominance in exploration tools leads to the perception that girls are less interested, which in turn influences future allocation decisions, widening the initial gap in experiential capital.

Figure 3: Gender gaps in access to devices

[3] Evidence from the US suggests that while girls in middle and high school often outperform boys academically, they also face greater mental health risks associated with excessive use of digital platforms. If GenAI follows a similar pattern, in which certain applications disproportionately affect specific groups, the focus should be on identifying which GenAI applications benefit students and which introduce risks, rather than simply increasing usage across the board.

[4] All charts in this section use a relative-gender-gap metric calculated as [](*Vp*/*Vt*)/(*Mp*/*Mt*)]−1, where *Vp* is the number of male participants in a given activity, *Vt* the total male population, *Mp* the number of female participants, and *Mt* the total female population. This formula expresses the gender gap as the relative difference in participation rates between male and female students. A negative value indicates a gap in favor of girls, while a positive value signals a gap in favor of boys. A value of zero denotes parity in relative terms. Figures 3–11 use a consistent color scale: violet represents a gap in favor of girls, whereas yellow represents a gap in favor of boys.



This process creates a powerful externality. As Acemoglu et al. (2024) argue in the context of digital platforms, algorithmic choices that maximize engagement can shift preferences and alter equilibria. A similar mechanism is at work here: When micro:bit kits are deployed without explicit, gender-sensitive guidance, an initial tendency for boys to engage first is often observed. Teachers may then (perhaps unconsciously) infer that girls are less interested, leading them to allocate future opportunities in ways that reinforce the initial imbalance. This creates a feedback loop (a negative externality) that widens the gap in experiential capital and perceived competence over time.

Therefore, equitable policy must look beyond nominal access figures and scrutinize the algorithmic and organizational processes that allocate learning opportunities. The goal is to identify and interrupt these negative externalities to ensure GenAI does not become the next micro:bit, a tool whose benefits are allocated unevenly from the start. The following sections analyze current platform usage through this lens, examining how existing patterns of engagement predict readiness for different GenAI functions.

3.3 Teacher Engagement and the Direction of Digital Tool Use

Understanding teacher readiness for GenAI requires mapping existing platform use to our functional taxonomy. Current Ceibal platforms, while not all generative, represent the existing trajectory of AI-enabled tool adoption, revealing how instructional tasks are already being automated and supported:

- CREA (learning management system with PowerBuddy) supports the preparation function. The recent addition of PowerBuddy, a GenAI assistant for drafting activities, represents a clear point of transition from nongenerative to generative AI, automating content-creation tasks.
- Matific and ALEKS (adaptive platforms) employ narrow AI algorithms for personalized learning paths, corresponding to reinforcement and feedback functions. They demonstrate teacher comfort with automating differentiated practice and assessment, a precursor to generative feedback systems.
- Robogarden (coding environment) uses rule-based AI tutors to support exploration and reinforcement through guided programming. It highlights engagement with open-ended problem-solving tasks, a key function for GenAI.

The analysis reveals that adoption is not neutral but is directionally biased, creating divergent pathways.

The most substantial gaps appear in exploration-focused activities. Male teachers are significantly more likely to request micro:bit boards and guide complex coding missions (Figure 3, right-hand-side panel). This suggests that the institutional and social processes for allocating resources for exploration are already creating externalities, favoring one group over another.

3.4. Student Pathways and the Accumulation of Digital Capital

Student usage patterns mirror and are reinforced by teacher practices, leading to an uneven accumulation of the skills and confidence (digital capital) needed to benefit from GenAI.

Structured Practice and Preparation

Girls demonstrate higher and more consistent engagement with platforms supporting practice and preparation. They show higher login rates, sustained use, and assignment completion on CREA (Figure 4), and they are more likely to complete both initial and final diagnostic tasks on adaptive platforms like Matific (Figure 5).

Figure 4: Platform usage data, CREA

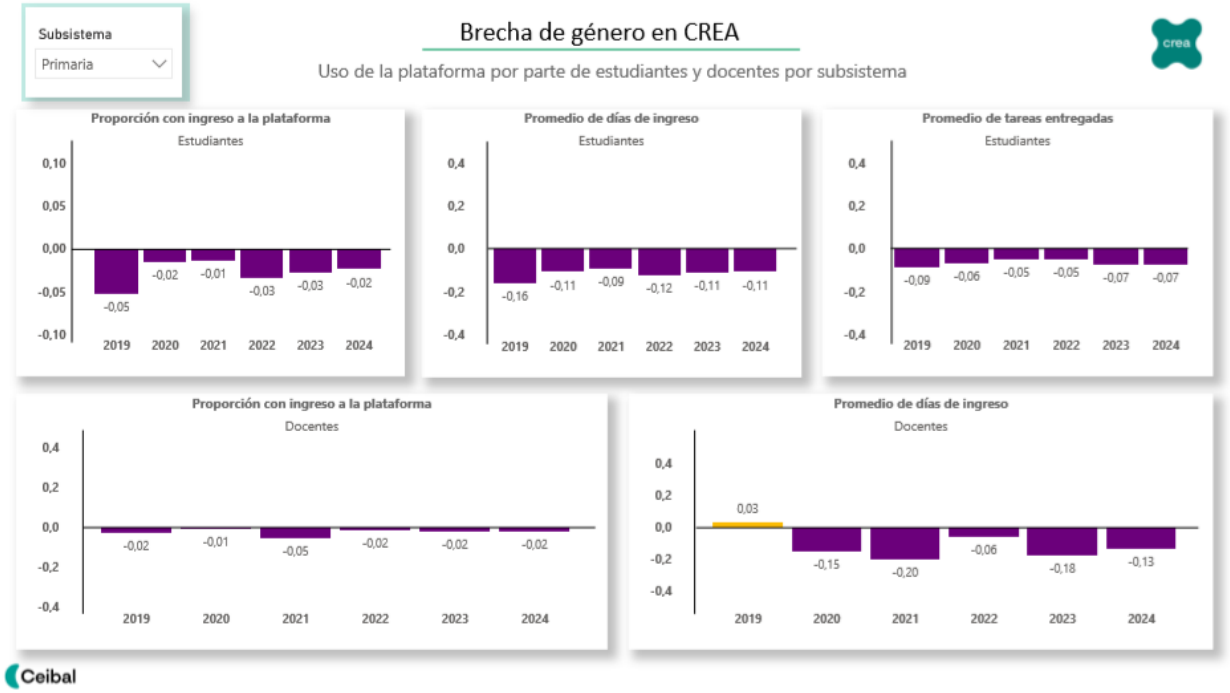
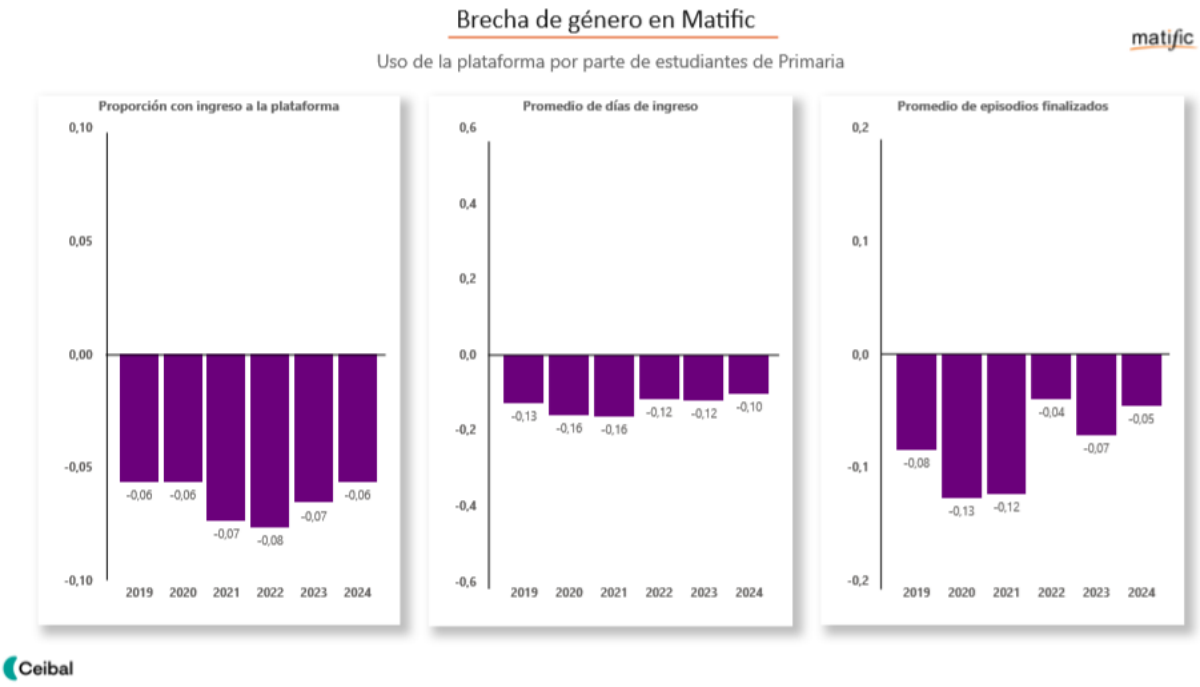


Figure 5: Platform usage data, Matific

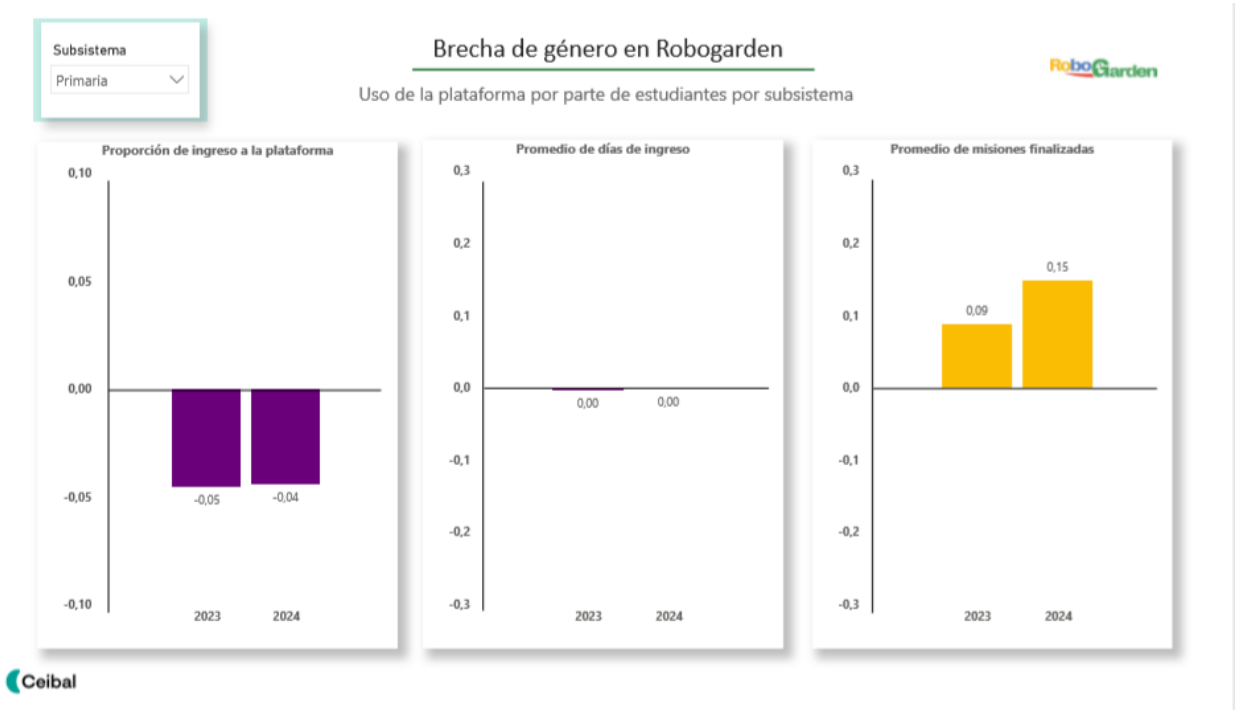


Open Exploration

A stark reversal occurs in platforms emphasizing the exploration function. While login rates for Robogarden are comparable, boys significantly outperform girls in mission completion by

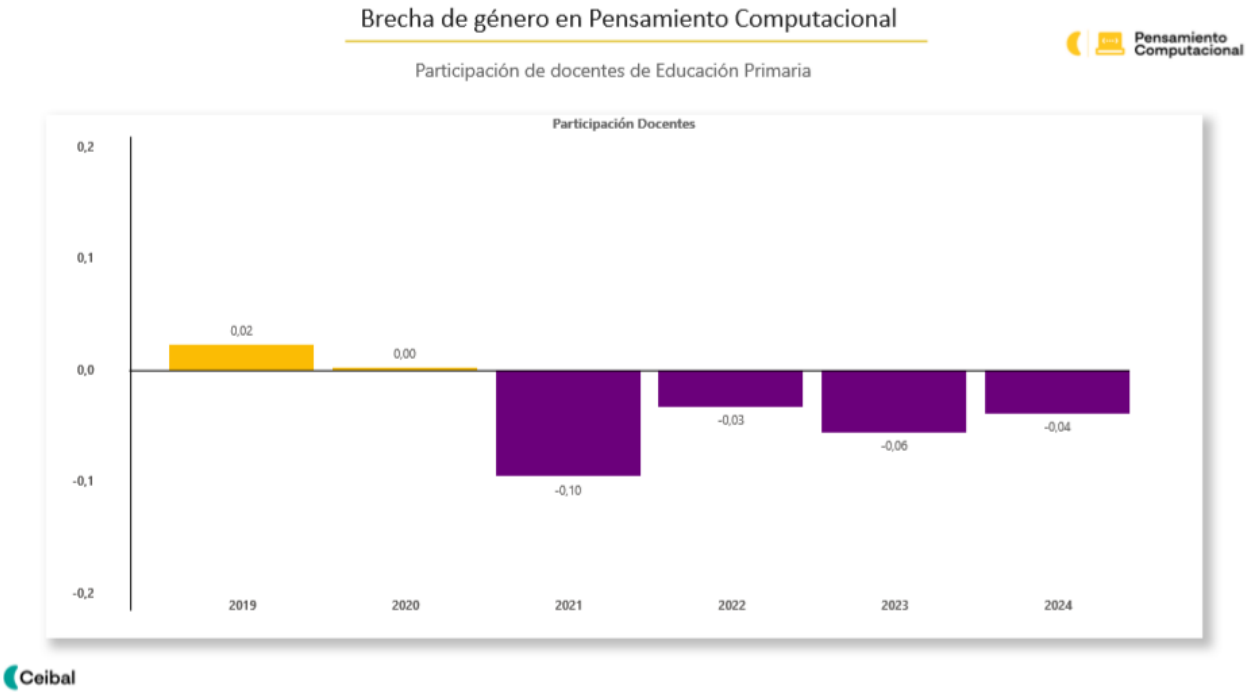
over 15 percentage points (Figure 6), indicating a substantial gap in persistence, confidence, and success in self-directed technical challenges.

Figure 6: Platform usage data, Robogarden



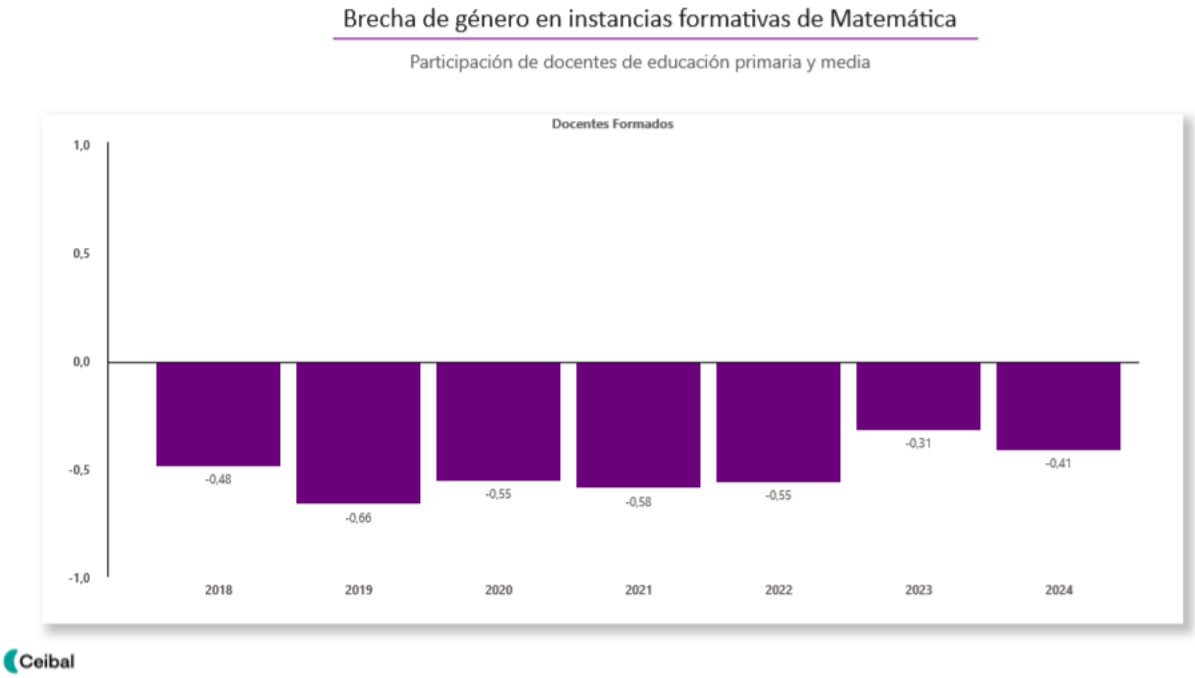
Female teachers show higher levels of participation in key areas related to digital and computational pedagogy (preparation function) by around four points (Figure 7).

Figure 7: Teacher-participation gender gap in Computational Thinking and AI Program, Ceibal



Data from ALEKS training (Bentancor, 2024) reveal a critical divergence in how reinforcement and exploration functions are used. Among trained math teachers, 71 percent of women use the platform mainly for drill and review (reinforcement), whereas 50 percent of men use it for knowledge construction (exploration). This is not merely a preference; it reflects how the same technology can be directed toward either routine practice (a form of automation) or conceptual exploration (a form of complementarity). As shown in Figure 8, women are overrepresented in professional development initiatives for mathematics by more than 40 percentage points. These patterns contradict common assumptions about lower female engagement with technology. They also suggest that when structured opportunities are available, women teachers take them up at high rates. However, unequal classroom conditions and differentiated support may still limit the translation of this participation into effective platform use.

Figure 8: Teacher gender gap in participation in math training and development courses, Ceibal

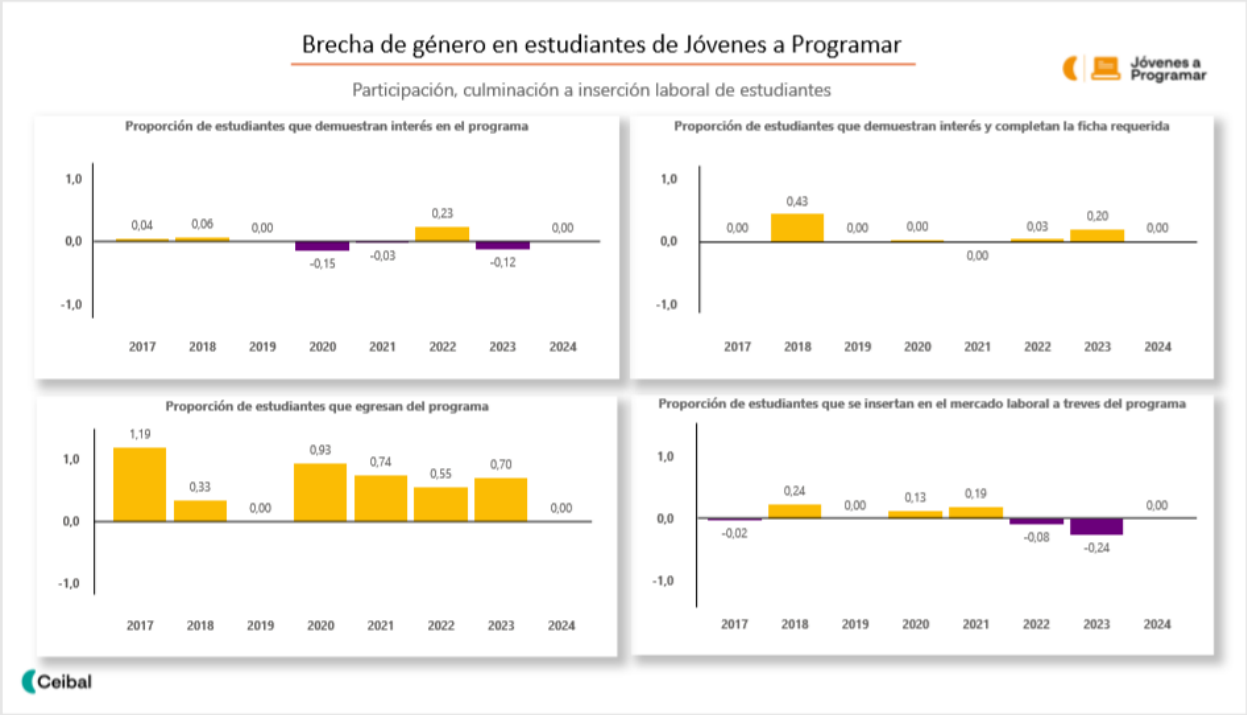


Student usage patterns mirror and are likely influenced by teacher practices, leading to an uneven accumulation of the skills and confidence (digital capital) needed to benefit from GenAI. Näslund-Hadley and Alonzo (2024) find that persistent gender gaps in learning outcomes across Latin America reflect both socialization patterns and unequal access to skill-building opportunities, especially in domains tied to the future of work.

The divergence extends into optional programs. Boys are overrepresented in competitive extracurricular programs like Jóvenes a Programar and the Video Games and Robotics Olympics (Figures 9, 10, and 11). Jóvenes a Programar trains young people aged 18 to 30 in programming, testing, English, and socioemotional skills to facilitate entry into the ICT sector. For example, a women-only version of the program demonstrated that in the absence of men, women improved their performance in areas like mathematics and logical reasoning (Gómez-Ruiz et al., 2024). These programs are not just learning opportunities; they are signals of belonging and competence in the technology domain. Their gendered participation creates a feedback loop: Success and visibility in these arenas reinforce the perception of boys as naturally more technically adept, which in turn influences teacher expectations and resource allocation (Porto et al., 2024). This is a clear example of the type of externality described by Acemoglu et al. (2024), in which early

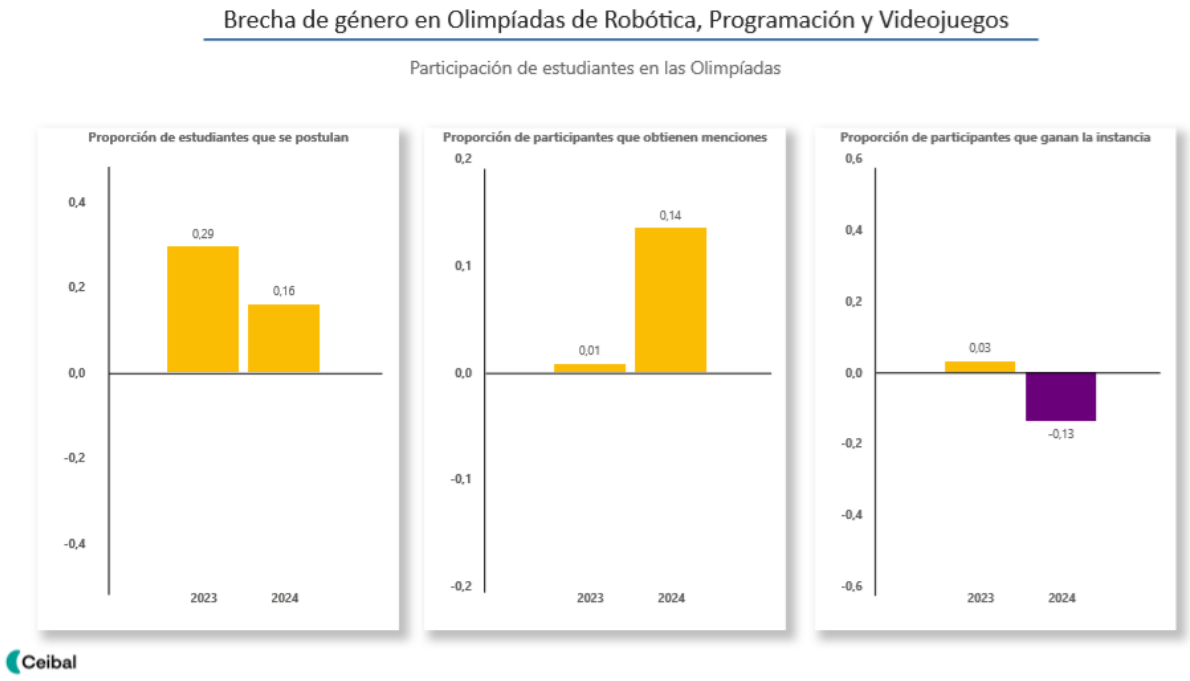
algorithmic or organizational choices (which programs are promoted, how they are designed) shift perceptions and participation, creating a new, less equitable equilibrium.

Figure 9: Gender gaps in participation, use, completion, and labor market insertion, Jóvenes a Programar



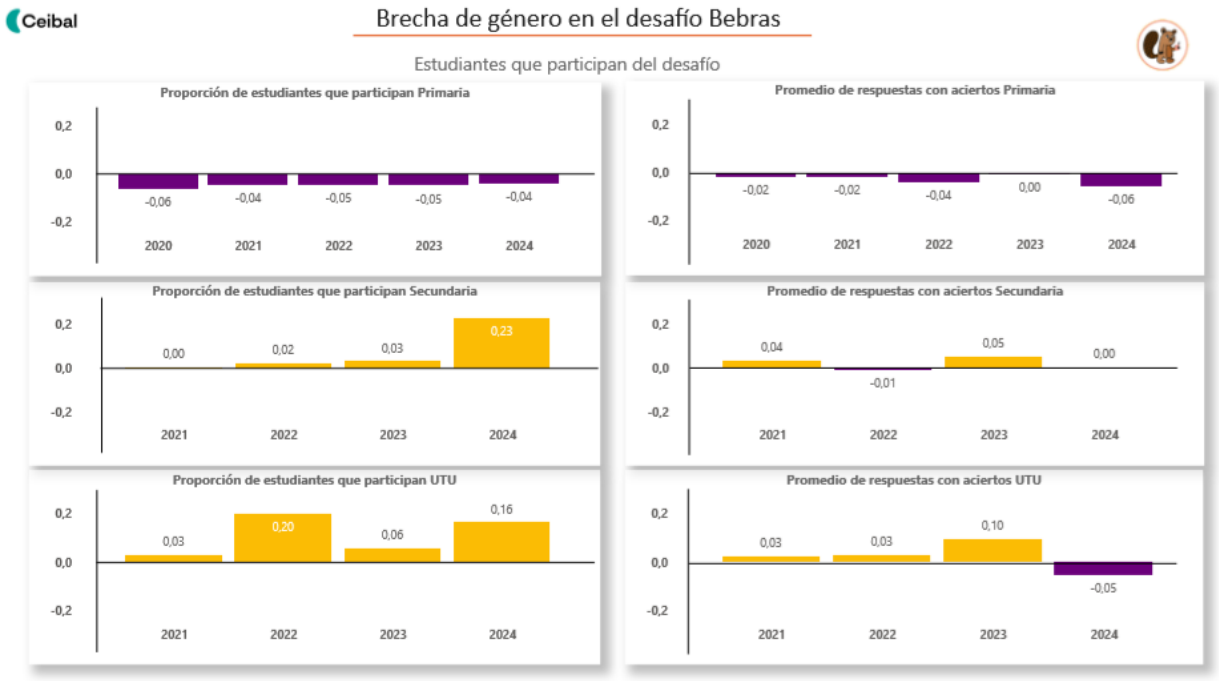
As Sulkunen (2013) argues, addressing literacy challenges among adolescents requires sustained, context-sensitive strategies that go beyond access and emphasize guided, purposeful engagement with texts and digital content. These extracurricular gaps matter for GenAI adoption because prior exposure to exploratory tasks conditions perceived usefulness. The pre-pilot therefore constrains exploration features and strengthens guided explanation and feedback in class.

Figure 10: Gender gaps in Ceibal Olympics of Robotics and Videogames



The gap is most pronounced in competitive or extracurricular programs. In the Bebras computational-thinking challenge (Figure 11), the gender split favors boys from the first cycle of secondary education onward. Even when girls participate, their completion and advancement rates are lower.

Figure 11: Gender gaps in Bebras Challenge



These patterns echo findings from other countries: When digital or STEM learning opportunities are optional, motivational and cultural factors often reproduce gendered engagement. In Uruguay, structural equity in device distribution has not fully translated into participatory equity in enrichment programs. While the integration of digital learning technologies has redefined early education in the country, it underscores the need to move from access to meaningful use and sustained engagement (Robano, 2024).

These gendered patterns of engagement matter for GenAI adoption because familiarity with digital problem-solving environments, such as Scratch, Python, or micro:bit, influences whether students see AI tools as relevant or usable. Girls may enter AI-rich learning environments with less prior exposure, not because of lower aptitude, but because of accumulated gaps in opportunity and support. While Uruguay has closed the access gap through Ceibal's universal device distribution, the next challenge is pedagogical: how to ensure that digital tools support learning in equitable, developmentally appropriate, and supervised ways. International preliminary studies suggest that teachers view GenAI tools like ChatGPT as useful but express concerns about classroom control, student dependency, and equity of access (Bada Basilio, 2024).

3.5 Connecting Platform Use to GenAI-Function Readiness: A Framework for Understanding Externalities

The current allocation of digital opportunities creates a predictable landscape for GenAI adoption:

- High Female-Readiness Functions: Preparation, reinforcement, and structured feedback. GenAI tools serving these functions are likely to see more equitable uptake, as they align with existing engagement patterns and strengths.
- High-Risk Functions (Male Dominated): Exploration and advanced explanation. These functions risk reproducing and exacerbating existing participation gaps. Their adoption is likely to be uneven without deliberate scaffolding and changes to the institutional processes that allocate them.
- Universal Functions: Basic explanation lacks clear gender patterns, suggesting potential for equitable adoption if designed neutrally.

This functional mapping explains why universal device access has failed to produce equitable engagement. The technology was directed toward specific uses, and the institutional context (teacher practices, program design, social norms) shaped its ultimate impact. This history provides an important lesson for GenAI: Its equity implications will be determined not by the technology itself, but by the power to direct its use. Will it be used to automate routine tasks for some while augmenting creative exploration for others? The current data suggests this is a default risk.

The functional taxonomy in Section 2 positions GenAI tools along classroom tasks, yet adoption depends on teacher characteristics that vary across the system. Professional tenure, subject area, level taught, and prior exposure to educational technology shape the likelihood that a teacher employs each GenAI function. Experience tells us that early-career teachers often value preparation aids that generate quizzes or lesson outlines, seasoned teachers more readily adopt exploration tools that invite open-ended projects, and STEM specialists frequently rely on adaptive feedback engines such as ALEKS. Humanities teachers, in contrast, may lean on explanation tools that provide step-by-step guidance. School socioeconomic context also matters: In settings in which attendance is uneven, motivation tools that deliver personalized nudges can compensate for lost seat time.

These patterns interact with the student gender gaps identified earlier. When teachers who guide robotics clubs receive micro:bit boards and possess the confidence to assign exploration tasks, students gain more opportunities to complete coding missions. Teachers who focus on reinforcement functions through Matific or ALEKS can reduce arithmetic anxiety for girls who already favor those platforms. The allocation of GenAI functions therefore mediates the

externalities described by Acemoglu et al. (2024): Algorithmic choices shift participation toward groups already positioned to benefit, and instructional practices adjust in response. Unequal teacher uptake can amplify or dampen gender imbalances even when hardware access appears balanced.

Linking the conceptual framework with the gender analysis thus requires attention to teacher profiles. Future evaluation should examine whether professional development that targets the underused functions (exploration for humanities teachers, motivation for low-attendance schools, and feedback analytics for novice staff) can align GenAI capabilities with equitable learning trajectories.

Box 1: Classroom vignette linking student experience, teacher mediation, and GenAI functions

A year 6 programming session with Scratch illustrates how teacher profiles and GenAI functions intersect with the gender patterns reported above. Eleven pupils shared a limited number of laptops, the classroom teacher was on site, and the computational-thinking specialist connected remotely. The task, "Build a drawing machine controlled by arrow keys," required pupils to save code to the sprite layer. Several learners (boys and girls alike) confused the stage and sprite, triggering repeated project failure, growing frustration, and eventual withdrawal. (In Scratch, a sprite is any on-screen object—such as a character, icon, or shape—that can be programmed with its own scripts and properties, as distinct from the background stage.)

Opportunity for GenAI. A feedback agent could deliver instant hints such as "save the block to the sprite, not the stage," and a reinforcement module could supply micro-challenges that rebuild the code step by step. These functions align with the preparation and explanation aids that novice teachers adopt first, as outlined in Section 3.4.

Role of the teacher. The on-site teacher noticed peer-to-peer scaffolding that the remote specialist could not see. Girls who hesitated to verbally request help leaned on nearby classmates, confirming the earlier finding that interpersonal support shapes platform engagement. The teacher's timely prompts restored participation, demonstrating the interpersonal layer that GenAI cannot replace.

Equity implications. When micro:bit boards or Scratch missions arrive without practical guidance, boys often take the first turn; the vignette shows how a teacher who

understands exploration tools can redistribute opportunities, while a conversational agent can prevent small procedural errors from snowballing into disengagement. The episode therefore connects the gender gaps in Section 3.2 with the teacher-function alignment in Section 3.4: Effective and equitable GenAI integration requires both adaptive technology and teachers who recognize when and how to activate each function.

3.6. Synthesis: The Default Trajectory for GenAI

The evidence from Ceibal's dashboard presents a clear and concerning diagnosis: Even in a system with equitable hardware access, the allocation of digital opportunities and functions is highly gendered, creating a default trajectory for GenAI adoption that threatens to amplify existing inequalities:

- Female students and teachers are on a pathway of high engagement with technology for preparation (for example, consistent use of CREA, completion of structured tasks on Matific) and reinforcement (for example, using ALEKS for drill and practice). This positions them as likely early adopters of GenAI tools that automate and enhance these specific functions.
- Male students and teachers are accumulating significantly more experiential capital and demonstrated success in the use of tools for exploration (for example, higher mission completion in Robogarden, dominance in micro:bit use, overrepresentation in competitive coding programs). This positions them to disproportionately benefit from GenAI's most disruptive and augmentative functions, which require comfort with open-ended, technical problem solving.

This functional divergence is not neutral. It is the product of institutional externalities (teacher practices, organizational routines, and program design) that systematically allocate opportunities and shape perceptions of technological affinity.

The critical implication for GenAI is that its equity impact will be determined not by the technology itself, but by the power to direct its use. The default trajectory, if left unaddressed, is clear: GenAI risks automating routine cognitive tasks for one group (primarily female) while augmenting creative and exploratory work for another (primarily male). This would cement a new, less equitable equilibrium, a classic case of technological adoption amplifying preexisting externalities (Acemoglu et al., 2024).

3.7. Bridging to Intervention and Recommendation

This diagnosis does not merely describe a problem; it provides a precise blueprint for action. It allows us to move from what is to what must be done by identifying the specific leverage points for intervention:

1. Targeting the Exploration Gap

   The pronounced male advantage in exploration-ready capital necessitates that GenAI tools designed for this function be deliberately scaffolded. Recommendations must focus on embedding supports (such as structured prompts, example-based learning, and failure-tolerant design) that lower the barrier to entry and interrupt the negative externality that discourages female participation.

2. Leveraging Preparation-Reinforcement Strength

   The high female readiness in preparation and reinforcement provides an important foundation. Recommendations should advocate leveraging these functions as an entry point to build digital confidence and competence among female students and teachers, creating a pathway that can later be extended to exploration.

3. Activating Teacher Mediation

   The divergent teacher practices reveal that professional development cannot be generic. Recommendations must be function specific and gender sensitive, training all teachers, especially those in humanities and early-career stages, to confidently deploy exploration tools and helping those focused on reinforcement to integrate more open-ended tasks.

4. Re-engineering Organizational Routines

   The gendered allocation of resources like micro:bit kits demands structural solutions. Recommendations must extend beyond the technology to include gender-aware implementation policies for how digital tools are promoted, distributed, and supported within schools.

Therefore, the experiment below is designed not to test a GenAI tool per se, but to test these specific mitigation mechanisms. Its findings will directly inform the recommendations in Section 5, providing empirical evidence useful for designing platforms, guiding teachers, and structuring policies to engineer a trajectory of equitable complementarity and ensure GenAI becomes a tool for reducing, rather than reinforcing, inequality.

4. Proposed Experimental Design

4.1. Intervention Philosophy and Pilot Phase

The diagnostic evidence presented in Section 3 reveals a system on a default trajectory toward inequitable adoption of GenAI, in which functions are likely to be allocated in ways that reinforce existing gendered externalities. Our experiment is designed to test whether a deliberately directed implementation can steer this technology toward a more complementary and equitable equilibrium.

To ground our formal RCT in real-world classroom dynamics, we first conducted an exploratory pre-pilot in June–August 2025. This pre-pilot, involving two sixth-grade classrooms, was not designed to test efficacy but to (1) refine the chatbot interface and functionality based on real student and teacher interactions, (2) finalize the teacher mediation protocol and guidance materials, and (3) validate our data-collection instruments (observation rubrics, survey questions) for the full-scale RCT. This iterative approach ensures the subsequent RCT tests a well-calibrated intervention, moving from diagnostic to pilot to rigorous evaluation.

4.2. RCT Design: Testing for Complementarity

The proposed RCT, to be deployed at scale in 2026, builds directly on insights from the pre-pilot and the diagnostic. Its primary objective is to causally identify whether the structured, teacher-mediated use of GenAI chatbots improves student learning outcomes and alters classroom interaction dynamics toward greater equity.

- Design. We will adopt a clustered randomization strategy, assigning entire schools to treatment and control conditions to prevent spillover effects.
- Intervention. The treatment group will use a chatbot, embedded in the familiar CREA platform. For students, the chatbot's capabilities will be restricted to the explanation function and the practice and feedback function. For teachers, the intervention will also include access to and training on GenAI tools for content preparation to enhance their lesson planning.
- Active Ingredient: Teacher Mediation. Crucially, the intervention is not the chatbot alone, but the chatbot plus a structured protocol for teacher mediation. Professional development will focus on empowering teachers to review, edit, and override GenAI suggestions, framing the tool as an assistant to their expertise.

4.3. Measuring What Matters: Complementarity and Equity

A key innovation of our design is that we define and measure success not only by final learning outcomes but by process metrics that directly capture complementarity and equity.

- Primary Outcomes:
    a. Student Learning Gains: Standardized computational-thinking assessment scores
    b. Measures of Complementarity
        - Teacher Mediation Logs: The frequency and type of teacher interventions (accept, edit, reject) on chatbot outputs.
        - Time Reallocation: From classroom observations, this is a measure of whether teachers in the treatment group spend less time on routine explanations and more time on high-quality, interactive instruction.
- Secondary Outcomes and Heterogeneity Analysis:
    a. Engagement Metrics: Usage intensity of the chatbot, broken down by function and, critically, by student gender.
    b. Equity Analysis: We will test for heterogeneous treatment effects across the dimensions identified in our diagnostic:
        - Student gender
        - Baseline student performance
        - Teacher gender and prior technological proficiency

This analysis is the ultimate test of whether our intervention can counteract, rather than reinforce, the preexisting externalities.

4.4. Analytical Approach and Limitations

We will use intent-to-treat analysis to estimate the causal impact of the intervention. The analysis will employ a regression model with baseline covariates to improve precision, and the heterogeneity analysis will be prespecified to rigorously test for differential impacts.

We acknowledge three key limitations. First, the outcomes may be sensitive to the quality of teacher training. Second, the study duration may be insufficient to capture long-term effects on student pathways and teacher skills. Nevertheless, this design provides an important test of whether directing GenAI toward complementarity is a viable policy option for mitigating educational inequality. A further limitation is the scope of available administrative data on teacher practices. While platform logs and mediation records provide valuable insights, they may not capture the full diversity of instructional choices across different teacher profiles. As a complementary strategy, future work could include structured surveys of preservice and in-service teachers to document

patterns of use, perceived barriers, and attitudes toward GenAI-enabled applications. This would enrich the quantitative analysis and strengthen the external validity of our findings.

5. Policy Implications: A Roadmap for Responsible Integration

To harness the promise of GenAI in education while addressing its risks, policymakers must adopt a strategy that is evidence based, inclusive, and pedagogically grounded. The following principles synthesize lessons from this study and ongoing debates on equitable technology integration.

5.1. Teacher Development

- Target Preparation-Function Gaps. Since female teachers show strong engagement with preparation-focused platforms, professional development should build on this strength while introducing advanced features that prevent overreliance on AI-generated content.
- Address Exploration-Function Disparities. Male teachers' higher adoption of exploration functions requires targeted interventions to ensure female teachers gain confidence with open-ended AI applications. Training should include mentorship, collaborative learning, and explicit bias recognition.
- Leverage Reinforcement-Function Strengths. Given female teachers' preference for structured practice platforms, GenAI tools serving reinforcement functions may achieve rapid adoption. Training should focus on pedagogical integration rather than basic functionality.

5.2 Student Pathway Development

- Build on Existing Engagement. Students already comfortable with structured platforms (CREA, Matific) may more readily adopt GenAI tools serving preparation and reinforcement functions. Curriculum design should create pathways from these familiar functions toward exploration applications.
- Address Exploration Gaps. The 15 percentage point gender gap in Robogarden completion suggests girls may disengage from open-ended technical challenges. GenAI exploration tools must include additional scaffolding, collaborative opportunities, and clear success criteria.

5.3 Platform Integration Strategy

- Leverage Existing Infrastructure. Embedding GenAI functions within trusted platforms like CREA ensures continuity and reduces adoption barriers. Teachers already comfortable

with preparation tasks in CREA may more readily accept AI-enhanced versions of these functions.

- Gradual Function Introduction. Rather than deploying all GenAI functions simultaneously, implementation should begin with high-adoption functions (preparation, reinforcement) before expanding to exploration and advanced explanation capabilities.

5.4 Evidence-Based Function Prioritization

Future experimentation should focus on functions that evidence suggests have the greatest impact:

1. High-Priority Functions. Explanation and reinforcement functions show consistent positive effects in international studies and align with existing teacher strengths in Uruguay.
2. Moderate-Priority Functions. Feedback functions require careful calibration to avoid teacher de-skilling while supporting formative assessment.
3. Caution-Required Functions. Exploration functions offer significant potential but require substantial scaffolding to prevent reproduction of existing gender gaps.

6. Conclusion and Next Steps

This paper advances a task-based framework to guide the responsible integration of GenAI in Latin American classrooms, with a particular focus on complementarity with teaching and on equity in access and use. Building on the empirical literature, fieldwork, and programmatic experience, it highlights the need to move beyond generalized enthusiasm or skepticism toward structured experimentation and inclusive design. The proposed RCT, informed by a pre-pilot in Ceibal Uruguay, aims to generate rigorous evidence on how chatbot-based AI tools interact with classroom dynamics and learning outcomes. The evidence from the Ceibal ecosystem demonstrates that equal device access does not automatically translate to equitable engagement with educational technology. Gender gaps emerge not in basic platform access but in how different groups engage with platforms serving different pedagogical functions. This finding has direct implications for GenAI policy: Simply providing access to AI tools will not ensure equitable outcomes unless implementation strategies account for differentiated function preferences and provide targeted support for underrepresented adoption patterns. Future research should continue to disentangle effects at the individual and systemic levels, with attention to the distributional consequences of AI use. Ultimately, the challenge for education systems in the region is not whether to adopt GenAI, but how to do so in ways that are pedagogically meaningful, socially inclusive, and institutionally grounded.

7. Research Agenda

While this paper provides a foundational framework for understanding GenAI adoption in education, several critical questions remain that will determine the success of policies concerning directed technological change. We identify 10 priority areas for future research that address the theoretical, methodological, and practical limitations of our current analysis.

Theoretical Development

Future research should develop formal models specifying the production functions through which teacher-AI complementarity generates learning outcomes and the precise mechanisms by which early technology adoption creates self-reinforcing participation patterns.

Causal Identification

Future studies should leverage exogenous variation in technology access or assignment to establish causal relationships between digital-tool adoption patterns and the gendered externalities we document.

Economic Analysis

Comprehensive cost-effectiveness analyses comparing GenAI-mediated instruction to alternative pedagogical interventions will be essential for scaling decisions and resource-allocation guidance.

System-Wide Effects

Future work should examine system-wide effects on teacher labor markets, resource allocation between schools, and shifts in educational demand that our partial equilibrium analysis cannot capture.

Distributional Analysis

Future research should analyze the distributional implications of GenAI implementation costs and their potential to exacerbate between-school inequalities using detailed budget and financing data.

Long-Term Impacts

Long-term follow-up studies will be necessary to capture dynamic effects on human capital formation and career trajectories that extend beyond our short-term experimental window.

Methodological Improvements

Future experimental designs should explore alternative randomization strategies that provide sufficient statistical power for detecting the gender-specific heterogeneous effects central to our theoretical framework.

Broader Outcomes

Subsequent studies should examine spillover effects on mathematics, literacy, and socioemotional skills to provide a comprehensive assessment of GenAI's educational impact beyond computational thinking.

External Validity

Replication studies in diverse institutional contexts will be needed to test the external validity of our findings beyond Uruguay's unique technological infrastructure and governance system.

Implementation Science

Future implementation research should systematically analyze political economy factors including teacher union preferences, parental attitudes, and bureaucratic capacity that determine scalability and sustainability. Addressing these research priorities will be essential for building the evidence base needed to guide responsible GenAI adoption in education systems across Latin America and beyond.

References

Acemoglu, Daron (2024) The Simple Macroeconomics of AI Daron Acemoglu Working Paper 32487

Acemoglu, Daron.2021. “Harms of AI.” National Bureau of Economic Research Working Paper 29247.

Acemoglu, D., and Johnson, S. (2023). Power and progress: our thousand-year struggle over technology and prosperity. First edition. New York, PublicAffairs.

Acemoglu, Daron, and Pascual Restrepo.2021. “Tasks, Automation, and the Rise in US Wage Inequality.” National Bureau of Economic Research Working Paper 28920.

Acemoglu, Daron, and Pascual Restrepo.2019. “Artificial Intelligence, Automation, and Work.” in The Economics of Artificial Intelligence: An Agenda., ed. Ajay Agrawal, Joshua Gans and Avi Goldfarb, 197–236. University of Chicago Press.

Acemoglu, Daron, and Pascual Restrepo.2018. “The Race between Man And Machine: Implications of Technology for Growth, Factor Shares, and Employment. “American Economic Review, 108(6): 1488–1542.

Aghion, Philippe, Benjamin F Jones, and Charles I Jones.2017. “Artificial Intelligence and Economic Growth.” National Bureau of Economic Research Working Paper 23928

Administración Nacional de Educación Pública [ANEP]. (2023). Uruguay en PISA 2022: Volumen 2. Calidad, equidad y metas educativas. Consejo Directivo Central, Dirección Sectorial de Planificación Educativa. ISBN: 978-9974-887-64-0. https://pisa.anep.edu.uy/sites/default/files/Recursos/Publicaciones/Informes/2022/Uruguay%20en%20PISA%202022_Volumen%202_Calidad,%20equidad%20y%20metas%20educativas.pdf

Arias Ortiz, E., Giambruno, C., Morduchowicz, A., and Pineda, B. (2024). El estado de la educación en América Latina y el Caribe 2023. https://doi.org/10.18235/0005515

Autor, David.2024. “Applying AI to Rebuild Middle Class Jobs.” National Bureau of Economic Research Working Paper 32140.

Ayres, R.U., and S. M. Miller.1983.Robotics, Applications and Social Implications. Ballinger Publishing Company

Bada Basilio, Manolito.2024. “Teachers’ Perspectives on the Integration of ChatGPT in Classroom Teaching and Learning.”

Bahroun, Z., Anane, C., Ahmed, V., and Zacca, A. (2023). Transforming Education: A Comprehensive Review of Generative Artificial Intelligence in Educational Settings through Bibliometric and Content Analysis. Sustainability, 15(17), 12983. https://doi.org/10.3390/su151712983

Bastani, Hamsa and Bastani, Osbert and Sungu, Alp and Ge, Haosen and Kabakcı, Özge and Mariman, Rei, Generative AI Can Harm Learning (July 15, 2024). The Wharton School Research Paper, Available at SSRN: https://ssrn.com/abstract=4895486 or http://dx.doi.org/10.2139/ssrn.4895486

Bentancor Gustavo (2025), “Evolución de la brecha de género en matemática en Uruguay: evidencia y reflexiones,” capítulo 3, página 56 en De la teoría a la práctica: construyendo equidad de género en STEM | Ceibal

Bhatt, M. P., Guryan, J., Khan, S. A., LaForest-Tucker, M., and Mishra, B. (2024). Can technology facilitate scale? Evidence from a randomized evaluation of high dosage tutoring (NBER Working Paper No. 32510). National Bureau of Economic Research. https://doi.org/10.3386/w32510

Bodonhelyi, Anna, Enkeleda Thaqi, Suleyman Ozdel, Efe Bozkir, and Enkelejda Kasneci.2024. “From Passive Watching to Active Learning: Empowering Proactive Participation in Digital Classrooms with AI Video Assistant. https://arxiv.org/html/2409.15843v1

Brynjolfsson, Erik, Danielle Li, and Lindsey R. Raymond. 2023. *Generative AI at Work*. Working Paper 31161, National Bureau of Economic Research. DOI: 10.3386/w31161. Revised November 2023.

Capdehourat German, Brian Lorenzo, Pedro Porteiro and Joaquín Trigo (2025) “On the effectiveness of LLMs for automatic grading of open-ended questions in Spanish” https://documentos.ceibal.edu.uy/portal/2025/02/URUCON_2024_camera_ready.pdf

Carvajal, Daniel, Catalina Franco, and Siri Isaksson.2024. “Will Artificial Intelligence Get in the Way of Achieving Gender Equality?” NHH Dept. of Economics 03. Revised in the Paper Series.

Castro, Juan F., Paul Glewwe, Alexandra Heredia-Mayo, Stephanie Majerowicz, Ricardo Montero (2024) Can Teaching Be Taught? Improving Teachers' Pedagogical Skills at Scale in Rural Peru1

Ceibal EduIA (2025) Publicaciones científicas—Ceibal

Ceibal AI Manifest (2025), available at: https://documentos.ceibal.edu.uy/portal/2025/01/Ceibal_Manifiesto-IA.pdf

Ceibal (2024a) Construyendo Inteligencia Artificial para la educación 2024 Contribuciones de María Eugenia Curi, Víctor Koleszar, Brian Lorenzo y Germán Capdehourat

Ceibal (2024b) Explorando experiencias de inteligencia artificial en el aula ISBN 978-9915-9704-5-5

Ceibal. (2023). Marco referencial para la enseñanza de la inteligencia artificial. Montevideo: Ceibal. ISBN: 9789915957845. Contribuciones de María Eugenia Curi, Víctor Koleszar, Brian Lorenzo y Germán Capdehourat.Marco referencial para la enseñanza de la inteligencia artificial

Curi, M. E., Koleszar, V., Capdehourat, G., Pereiro, E., Lorenzo, B., and Folgar, L. (2020). Construyendo inteligencia artificial para la educación. Cristobal Cobo (Prólogo). ISBN 978-9915-9660-0-7.

Deng Ruiqi Deng, Maoli Jiang, Xinlu Yu, Yuyan Lu, Shasha Liu, (2025) Does ChatGPT enhance student learning? A systematic review and meta-analysis of experimental studies, Computers and Education, Volume 227, 2025, 105224, ISSN 0360-1315, https://doi.org/10.1016/j.compedu.2024.105224 (https://www.sciencedirect.com/science/article/pii/S0360131524002380)

ECLAC. (2024). Prevención y reducción del abandono escolar en América Latina y el Caribe. CEPAL.https://repositorio.cepal.org/server/api/core/bitstreams/b80b4d28-e4df-4c5c-9659-dd78d4d1b527/content

Eisfeldt, Andrea L., and Gregor Schubert.2024. "AI and Finance." National Bureau of Economic Research Working Paper 33076.

Fodouop Kouam, Arthur William and Muchowe, Regis. (2024). Exploring graduate students' perception and adoption of AI chatbots in Zimbabwe: Balancing pedagogical innovation and development of higher-order cognitive skills. Journal of Applied Learning and Teaching. 7. 1–11. 10.37074/jalt.2024.7.1.12.

Gerlich, M. (2025). AI Tools in Society: Impacts on Cognitive Offloading and the Future of Critical Thinking. Societies, 15(1), 6. https://doi.org/10.3390/soc15010006

Goh E, Gallo R, Hom J, Strong E, Weng Y, Kerman H, Cool J, Kanjee Z, Parsons AS, Ahuja N, Horvitz E, Yang D, Milstein A, Olson APJ, Rodman A, Chen J H. Influence of a Large Language Model on Diagnostic Reasoning: A Randomized Clinical Vignette Study. medRxiv [Preprint]. 2024 Mar 14:2024.03.12.24303785. doi: 10.1101/2024.03.12.24303785. Update in: JAMA Netw Open. 2024 Oct 1;7(10):e2440969. doi: 10.1001/jamanetworkopen.2024.40969. PMID: 38559045; PMCID: PMC10980135.

Henkel, O., Horne-Robinson, H., Kozhakhmetova, N., Lee, A. (2024). Effective and Scalable Math Support: Experimental Evidence on the Impact of an AI-Math Tutor in Ghana. In: Olney, A.M., Chounta, IA., Liu, Z., Santos, O.C., Bittencourt, I. I. (eds) Artificial Intelligence in Education. Posters and Late Breaking Results, Workshops and Tutorials, Industry and Innovation Tracks, Practitioners, Doctoral Consortium and Blue Sky. AIED 2024. Communications in Computer and Information Science, vol 2150. Springer, Cham. https://doi.org/10.1007/978-3-031-64315-6_34

Holmes, W; Bialik, M; Fadel, C; (2019) Artificial Intelligence in Education Promises and Implications for Teaching and Learning. (1st ed.). Center for Curriculum Redesign: MA, USA.

Hoskins, Bryony Louise, and Ulf Fredriksson.2008.Learning to Learn: What is it and Can it Be Measured. EUR 23432 EN, Luxembourg (Luxembourg):OPOCE. JRC46532.

INEEd (2024). Reporte del Mirador Educativo 10. Ausentismo crónico en educación primaria pública: caracterización del período 2019–2023. Recuperado de https://www.ineed.edu.uy/images/Mirador/Reportes/Ausentismo-cronico-en-educacion-primaria-publica-2019-2023.pdf

INTEF (2022). Pushing the frontiers of education with intelligent technologies. Available at: https://intef.es/wp-content/uploads/2022/10/2022_01_Pushing-the-frontiers_INTEF.pdf.

Joubin, Alexa Alice (2024). “Enhancing the Trustworthiness of Generative Artificial Intelligence in Responsive Pedagogy in the Context of Humanities Higher Education.”General Aspects of Applying Generative AI in Higher Education: Opportunities and Challenges,

ed. Mohamed Lahby, Yassine Maleh, Antonio Bucchiarone and Satu Elisa Schaeffer, 207–220. Cham: Springer Nature Switzerland.

Kim, K., and Kwon, K. (2024). Designing an Inclusive Artificial Intelligence (AI) Curriculum for Elementary Students to Address Gender Differences With Collaborative and Tangible Approaches. Journal of Educational Computing Research, 62(7), 1617–1644. https://doi.org/10.1177/07356331241271059 (Original work published 2024)

Lehmann, Matthias and Cornelius, Philipp B. and Sting, Fabian J., AI Meets the Classroom: When Do Large Language Models Harm Learning? (March 07, 2025). Available at SSRN: https://ssrn.com/abstract=4941259 or http://dx.doi.org/10.2139/ssrn.4941259

Luckin, Rose; Holmes, Wayne; Griffiths, Mark and Forcier, Laurie B. (2016). Intelligence Unleashed. An argument for AI in Education Open Ideas; Pearson Education, London.

Malik, Rizwaan, Dorna Abdi, Rose Wang, and Dorottya Demszky. (2025). Scaffolding Middle-School Mathematics Curricula With Large Language Models. (EdWorkingPaper: 24–1028). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/b47y-mh41

Mollick, Ethan R. and Mollick, Lilach, Assigning AI: Seven Approaches for Students, with Prompts (September 23, 2023). The Wharton School Research Paper, Available at SSRN: https://ssrn.com/abstract=4475995 or http://dx.doi.org/10.2139/ssrn.4475995

OECD (2023), Empowering Young Children in the Digital Age, Starting Strong, OECD Publishing, Paris, https://doi.org/10.1787/50967622-en .

OECD. (2024). Equity in education and on the labour market: Main findings from Education at a Glance 2024 (OECD Education Policy Perspectives No. 107). OECD Publishing. https://doi.org/10.1787/b502b9a6-en

Pappagallo, S. (2024), Chatbots in Education: A Dual Perspective on Innovation and Ethics—Journal of Digital Pedagogy

Porto, C., Pereiro, E., Curi, M. E., Koleszar, V., and Urruticoechea, A. (2024). Gender perspective in the computational thinking program of Uruguay: teachers' perceptions and results of the Bebras tasks. Journal of Research on Technology in Education, 1–15.

Robano, V., I. González and O. Galland (2025) "De la teoría a la práctica: construyendo equidad de género en el ecosistema STEM de Ceibal"

https://documentos.ceibal.edu.uy/portal/2025/04/De-la-teoria-a-la-practica_construyendo-equidad-de-genero-en-STEM.pdf

Robano, V. (2024) Tecnologías Digitales para la Educación, https://ceibal-fichas-edtech.netlify.app/

Robano, V. (2023) Equidad de género en el ecosistema STEM de Ceibal ISBN: 978-9915-9470-5-1 Equidad de género en el ecosistema STEM de Ceibal https://documentos.ceibal.edu.uy/portal/2023/04/Equidad_Genero_STEM_Ceibal.pdf

Saavedra, Jaime, Cristóbal Cobo, Emanuela Di Gropello, Ezequiel Molina, Helena Rovnerr, Alex Twinomugisha, (2024), Educating for the (present and) future: using Artificial Intelligence (AI) to address the learning crisis

Sahito, Z. H., Sahito, F. Z., and Imran, M. (2024). The role of artificial intelligence (AI) in personalized learning: A case study in K-12 education. *Global Educational Studies Research Journal*, 9(3), 15. http://dx.doi.org/10.31703/gesr.2024(IX-III).15

San Román Inzaurralde, N. (2023) Matemática temprana y género : un análisis de resultados en escuelas uruguayas [en línea] Trabajo final de grado. Montevideo: Udelar. FP https://hdl.handle.net/20.500.12008/37071

Sethi, Rajiv, and Rohini Somanathan (2023) "Meritocracy and Representation. Journal of Economic Literature, 61(3): 941–57.

Son, T. (2024). Intelligent Tutoring Systems in Mathematics Education: A Systematic Literature Review Using the Substitution, Augmentation, Modification, Redefinition Model. Computers, 13(10), 270.

Su, Jiahong, and Weipeng Yang.2023. "Unlocking the Power of ChatGPT: A Framework for Applying Generative AI in Education." ECNU Review of Education, 6(3): 355–366.

Sulkunen, S. (2013). Adolescent Literacy in Europe: An Urgent Call for Action. European Journal of Education, 48(4), 528–542.

Varsik, S., and L. Vosberg. 2024. "The Potential Impact of Artificial Intelligence on Equity and Inclusion in Education." OECD Artificial Intelligence Papers, (23).

Vanzo, A., Pal Chowdhury, S., and Sachan, M. (2024). GPT-4 as a Homework Tutor can Improve Student Engagement and Learning Outcomes (arXiv preprint No. 2409.15981). arXiv. https://doi.org/10.48550/arXiv.2409.15981

Yeyati, Eduardo Levy (2024) Automatizados: Vida y trabajo en tiempos de inteligencia artificial ISBN 978-950-49-8631-7

Wang, F., Cheung, A. C. K., Neitzel, A. J., and Chai, C. S. (2024a). Does Chatting with Chatbots Improve Language Learning Performance? A Meta-Analysis of Chatbot-Assisted Language Learning. Review of Educational Research, 0(0). https://doi.org/10.3102/00346543241255621

Wang, R. E., Ribeiro, A. T., Robinson, C. D., Loeb, S., and Demszky, D. (2024b). TutorCopilot: A large language model-based tool for supporting high-quality tutoring [Working paper]. Stanford University. https://arxiv.org/abs/2410.03017

Zhai, C., Wibowo, S. and Li, L. D. The effects of over-reliance on AI dialogue systems on students' cognitive abilities: a systematic review. Smart Learn. Environ. 11, 28 (2024). https://doi.org/10.1186/s40561-024-00316-7

Glossary

Term	Definition
AI-tutor	An AI-powered tool designed to provide tailored educational assistance to students
Chatbot	A computer program designed to simulate conversation with a human user, often used to provide information or assistance. Powered by generative AI, chatbots like ChatGPT offer dynamic, natural conversations.
ChatGPT	A specific implementation of a generative AI chatbot, developed by OpenAI, based on a GPT (Generative Pretrained Transformer) model. It excels in generating humanlike, contextually relevant text. ChatGPT combines the general capabilities of generative AI and the conversational focus of chatbots.
CREA (Schoology)	A digital learning platform used in Uruguay's public education system
CTP	Computational-Thinking Program
Generative AI	A type of artificial intelligence that creates new content, such as text, images, audio, or video, based on input data or prompts. An umbrella term encompassing various models and tools, including LLMs, chatbots, and ChatGPT, that generate outputs rather than just analyzing or categorizing data.
LLM	A computer program that uses very large collections of language data in order to understand and produce text in a way that is similar to the way humans do large language model noun— Definition, pictures, pronunciation and usage notes | Oxford Advanced Learner's Dictionary at OxfordLearnersDictionarie A subset of generative AI. LLMs are the foundation for building tools like Chatbots and ChatGPT, enabling them to understand context and generate text-based outputs.
STEM	STEM is an acronym for Science, Technology, Engineering and Mathematics